

Kapitel 9: Verfahren für Nominaldaten

Eindimensionaler Chi²-Test

Berechnen der Effektgröße w^2

Die empirische Effektgröße w^2 gibt die Größe der Abweichung zwischen den erwarteten und beobachteten Häufigkeiten in standardisierter Form an. Das Maß w^2 ist inhaltlich schwer zu interpretieren. Nach der Klassifikation von Cohen (1988) gelten jedoch die folgenden Konventionen (vgl. Kap. 9.1.4):

w^2 bis 0,01 ; w bis 0,1	kleiner Effekt
w^2 bis 0,09 ; w bis 0,3	mittlerer Effekt
w^2 bis 0,25 ; w bis 0,5	großer Effekt

Diese Werte können als Anhaltspunkt verwendet werden, um eine Stichprobenumfangsplanung bzw. eine Teststärkeanalyse a priori durchzuführen. Alternativ kann die erwartete Effektstärke aus vorangegangenen Untersuchungen geschätzt werden. Das Maß w^2 relativiert die absolute Größe der Abweichungen (Chi²-Wert) an der Stichprobengröße: $\hat{w}^2 = \frac{\chi^2}{N}$.

Durch diese Standardisierung lassen sich die Effektgrößen auch aus verschiedenen Untersuchungen miteinander vergleichen. Obige Formel wird auch benötigt, um die Teststärke für eine Untersuchung a posteriori zu schätzen.

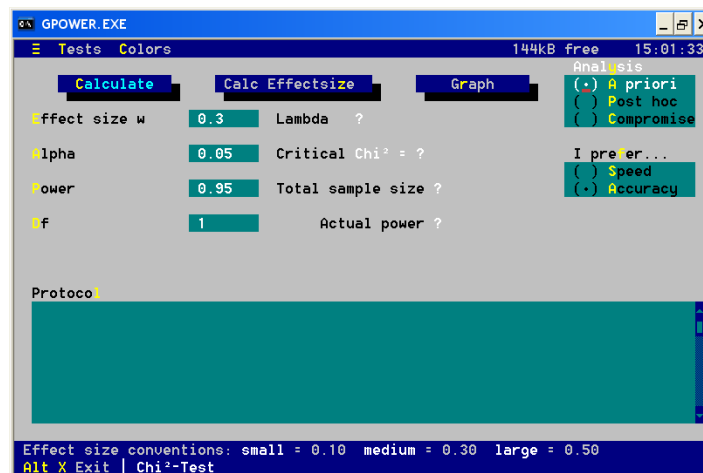
Alternativ kann in GPower unter „Tests“ → „Chi²-Tests“ → „Calc Effectsize“ die Effektstärke mit Hilfe von Zellwahrscheinlichkeiten berechnet werden, die unter der Null- und Alternativhypothese erwartet werden. Wir beschränken uns an dieser Stelle auf die Berechnung der empirischen Effektstärke wie oben beschrieben.

Stichprobenumfangsplanung bzw. Teststärkeanalyse a priori

Um zu ermitteln, wie viele Versuchspersonen benötigt werden, um einen erwarteten Effekt mit einer bestimmten Power von entdecken, muss eine Teststärkeanalyse durchgeführt werden (siehe Kap. 9.1.6). Öffnen Sie hierzu GPower und wählen Sie den Menüpunkt „Tests“ → „Chi²-Tests“. Sie sehen nun folgendes Fenster vor sich:

GPower-Ergänzungen

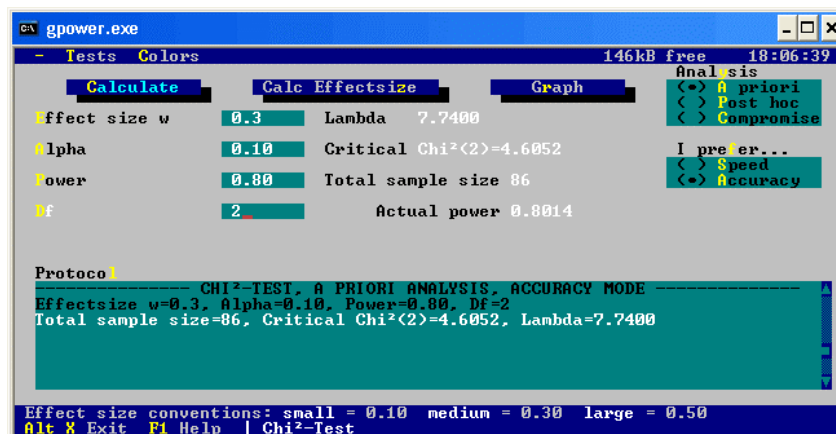
Rasch, Frieze, Hofmann & Naumann (2006). *Quantitative Methoden. Band 2* (2. Auflage). Heidelberg: Springer.



Achten Sie darauf, dass im Menü rechts (Analysis) die Option „A priori“ markiert ist. Tragen Sie die empirische Effektgröße w , die Wurzel (!) aus w^2 , unter „Effect size w “ ein und spezifizieren Sie das α -Fehlerniveau, sowie die gewünschte Teststärke (Power). Die Freiheitsgrade für den eindimensionalen χ^2 -Test ergeben sich nach der Berechnungsvorschrift: $df = k - 1$, wobei k der Anzahl der untersuchten Gruppen oder Kategorien entspricht (Kap. 9.1.2).

Wie viele Probanden müssen erhoben werden, um in einem eindimensionalen χ^2 -Test mit drei Gruppen ($df = 2$) einen mittleren Effekt der Größe $w^2 = 0,09$, also $w = 0,3$ mit einer Wahrscheinlichkeit von 80%, bei einem α -Fehlerniveau von 10% zu entdecken?

Wenn wir die entsprechenden Eingaben machen, liefert das Programm GPower eine Stichprobengröße von $N = 86$:



Berechnen der Teststärke a posteriori

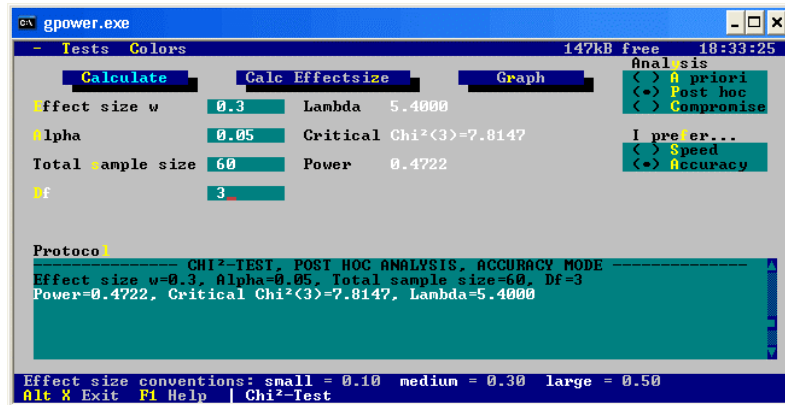
Um die Teststärke einer Studie a posteriori zu schätzen, öffnen Sie GPower und wählen Sie den Menüpunkt „Tests“ → „Chi²-Tests“ und markieren die Option „Analysis“ → „Post hoc“. Geben Sie dazu den empirisch ermittelten Effekt w , das α -Niveau, die tatsächliche Stichprobengröße sowie die Freiheitsgrade des Tests ein. GPower ermittelt dann die Power des Tests, also die Wahrscheinlichkeit, mit der Sie unter den bestehenden Bedingungen einen Effekt der bestehenden Größe auch tatsächlich entdecken konnten.

Quelle: <http://www.quantitative-methoden.de>

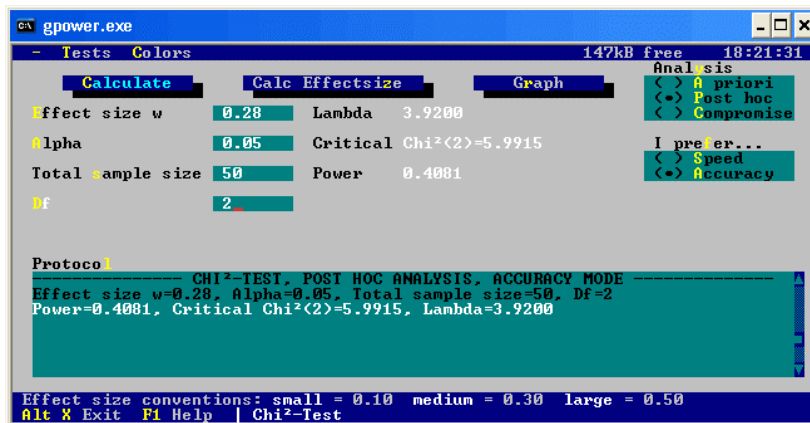
GPower-Ergänzungen

Rasch, Frieze, Hofmann & Naumann (2006). *Quantitative Methoden. Band 2* (2. Auflage). Heidelberg: Springer.

Ein Forscher hat über seine vier Bedingungen hinweg ein nicht signifikantes Ergebnis erzielt ($\alpha = 5\%$). In jeder Bedingung erfasste er 15 Versuchspersonen. Der Forscher erachtet einen mittleren Effekt von $w = .30$ für inhaltlich relevant. Wie groß war seine Chance, einen Effekt dieser Größe zu finden? Die Berechnung mit GPower zeigt, dass der Forscher eben so gut eine Münze hätte werfen können.



Ein anderer Forscher möchte die Effektstärke für seinen empirisch ermittelten Effekt berechnen. Nehmen wir beispielsweise an, er habe bei 50 Versuchspersonen und zwei Freiheitsgraden einen empirischen χ^2 -Wert von 3,90 ermittelt, der bei einem Signifikanzniveau von $\alpha = 5\%$ nicht signifikant geworden ist ($p = .14$). Unter Berücksichtigung der entsprechenden Effektgröße $w = .28$ und der restlichen Daten, erhalten Sie folgenden Output:



Neben dem kritischen χ^2 -Wert (5,99), der die Signifikanzgrenze markiert, ermittelt GPower eine Post hoc Teststärke von .4081. Das heißt, die Wahrscheinlichkeit, den bestehenden Effekt mit der gegebenen Stichprobengröße als signifikant nachzuweisen, war sehr gering. Insofern ist die Schlussfolgerung, es bestehe kein Effekt, in diesem Fall nicht gerechtfertigt. Vielmehr ist die richtige Schlussfolgerung, dass der Effekt zwar von mittlerer Größe ist, aber eben die Power nicht ausreichend hoch war, um ihn als statistisch signifikant nachzuweisen.

Der zweidimensionale und der Vierfelder Chi²-Test

Beim zweidimensionalen χ^2 -Test aus Kapitel 9.2 sowie beim Vierfelder Chi²-Test (Kap. 9.3) ist die Vorgehensweise bei der Teststärkenanalyse ähnlich zum eindimensionalen Fall. Auch hier

wird die Effektstärke über die Formel $\hat{w}^2 = \frac{\chi^2}{N}$ bestimmt, und es gelten dieselben Konventionen

für die Effektgröße wie oben vorgestellt. Lediglich die Berechnung der Freiheitsgrade erfolgt nun nach der Beziehung $df = (k - 1) \cdot (l - 1)$, wobei k und l die Anzahl der Gruppen bzw. Kategorien für die beiden untersuchten Dimensionen angibt (beim Vierfelder Chi²-Test betragen die Freiheitsgrade immer 1). Da die Vorgehensweise ansonsten identisch zum eindimensionalen Fall ist, wird hier auf eine Beispielrechnung verzichtet. Allerdings ist noch darauf hinzuweisen, dass GPower nicht mit den alternativen Effektstärken Cramers Index (Kap. 9.2.3) und dem Phi-Koeffizienten (Kap. 9.3.1) arbeitet, alle Berechnungen erfolgen über w .